

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## (i) APPLICANT:

(A) NAME: Rijksuniversiteit Leiden  
(B) STREET: Stationsweg 46  
(C) CITY: Leiden  
(D) STATE: Zuid-Holland  
(E) COUNTRY: the Netherlands  
(F) POSTAL CODE (ZIP): 2312 AV

(A) NAME: Nederlandse Organisatie voor Wetenschappelijk  
Onderzoek/Chemische Wetenschappen/STW  
(B) STREET: Laan van Nieuw Oost Indie 131  
(C) CITY: The Hague  
(D) STATE: Zuid-Holland  
(E) COUNTRY: the Netherlands  
(F) POSTAL CODE (ZIP): 2593 BM

(ii) TITLE OF INVENTION: Reducing branching and enhancing  
fragmentation in culturing filamentous microorganisms.

(iii) NUMBER OF SEQUENCES: 13

## (iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

## (v) CURRENT APPLICATION DATA:

PRIORITY APPLICATION NUMBER : EP 98202148.7

## (2) INFORMATION FOR SEQ ID NO: 1:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 438 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) Original Source: (A) Organism: Streptomyces griseus  
(B) Strain: ATTC 23345

## (ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 1..408  
(D) OTHER INFORMATION: /product= "SsgA"  
/gene= "ssgA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATGCGCGAGT CGGTTCAAGC AGAGGTCATG ATGAGCTTCC TCGTCTCCGA GGAGCTCTCG  
60

TTCCGTATTC CGGTGGAGCT CCGATACGAG GTCGGCGATC CGTATGCCAT CCGGATGACG  
120

TTCCACCTTC CCGGCGATGC CCCTGTGACC TGGGCGTTTC GCCGCGAGCT GCTGCTGGAC  
180

GGGCTCAACA GCCCGAGCGG CGACGGCGAT GTGCACATCG GCCCGACCGA GCCCGAGGGC  
240

CTCGGAGATG TCCACATCCG GCTCCAGGTC GGCGCGGACC GTGCGCTGTT CCGGGCGGGG  
300

ACGGCACCGC TGGTGGCGTT CCTCGACCGG ACGGACAAGC TCGTGCCGCT CGGCCAGGAG  
360

CACACGCTGG GTGACTTCGA CGGCAACCTG GAGGACGCAC TGGGCCCGCAT CCTCGCCGAG  
420

GAGCAGAACG CCGGCTGA  
438

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 408 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptomyces griseus*
- (B) STRAIN: ATTC 23345

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..408
- (D) OTHER INFORMATION: /product= "SsgA"  
/gene= "ssgA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

ATG AGC TTC CTC GTC TCC GAG GAG CTC TCG TTC CGT ATT CCG GTG GAG  
48  
Met Ser Phe Leu Val Ser Glu Glu Leu Ser Phe Arg Ile Pro Val Glu  
1 5 10 15

CTC CGA TAC GAG GTC GGC GAT CCG TAT GCC ATC CGG ATG ACG TTC CAC  
96  
Leu Arg Tyr Glu Val Gly Asp Pro Tyr Ala Ile Arg Met Thr Phe His  
20 25 30

CTT CCC GGC GAT GCC CCT GTG ACC TGG GCG TTC GGC CGC GAG CTG CTG  
144  
Leu Pro Gly Asp Ala Pro Val Thr Trp Ala Phe Gly Arg Glu Leu Leu  
35 40 45

CTG GAC GGG CTC AAC AGC CCG AGC GGC GAC GGC GAT GTG CAC ATC GGC  
192  
Leu Asp Gly Leu Asn Ser Pro Ser Gly Asp Gly Asp Val His Ile Gly  
50 55 60

CCG ACC GAG CCC GAG GGC CTC GGA GAT GTC CAC ATC CGG CTC CAG GTC

240  
 Pro Thr Glu Pro Glu Gly Leu Gly Asp Val His Ile Arg Leu Gln Val  
 65 70 75 80

GGC GCG GAC CGT GCG CTG TTC CGG GCG GGG ACG GCA CCG CTG GTG GCG  
 288  
 Gly Ala Asp Arg Ala Leu Phe Arg Ala Gly Thr Ala Pro Leu Val Ala  
 85 90 95

TTC CTC GAC CGG ACG GAC AAG CTC GTG CCG CTC GGC CAG GAG CAC ACG  
 336  
 Phe Leu Asp Arg Thr Asp Lys Leu Val Pro Leu Gly Gln Glu His Thr  
 100 105 110

CTG GGT GAC TTC GAC GGC AAC CTG GAG GAC GCA CTG GGC CGC ATC CTC  
 384  
 Leu Gly Asp Phe Asp Gly Asn Leu Glu Asp Ala Leu Gly Arg Ile Leu  
 115 120 125

GCC GAG GAG CAG AAC GCC GGC TG  
 408  
 Ala Glu Glu Gln Asn Ala Gly  
 130 135

## (2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 135 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Met Ser Phe Leu Val Ser Glu Glu Leu Ser Phe Arg Ile Pro Val Glu  
 1 5 10 15

Leu Arg Tyr Glu Val Gly Asp Pro Tyr Ala Ile Arg Met Thr Phe His  
 20 25 30

Leu Pro Gly Asp Ala Pro Val Thr Trp Ala Phe Gly Arg Glu Leu Leu  
 35 40 45

Leu Asp Gly Leu Asn Ser Pro Ser Gly Asp Gly Asp Val His Ile Gly  
 50 55 60

Pro Thr Glu Pro Glu Gly Leu Gly Asp Val His Ile Arg Leu Gln Val  
 65 70 75 80

Gly Ala Asp Arg Ala Leu Phe Arg Ala Gly Thr Ala Pro Leu Val Ala  
 85 90 95

Phe Leu Asp Arg Thr Asp Lys Leu Val Pro Leu Gly Gln Glu His Thr  
 100 105 110

Leu Gly Asp Phe Asp Gly Asn Leu Glu Asp Ala Leu Gly Arg Ile Leu  
 115 120 125

Ala Glu Glu Gln Asn Ala Gly  
 130 135

## (2) INFORMATION FOR SEQ ID NO: 4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 408 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Streptomyces albus G
- (B) STRAIN: ATCC 3004

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..408
- (D) OTHER INFORMATION: /product= "SsgA"  
/gene= "ssgA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```

ATG AGC TTC CTC GTC TCC GAG GAG CTC GCC TTC CGC ATC CCG GTG GAG
48
Met Ser Phe Leu Val Ser Glu Glu Leu Ala Phe Arg Ile Pro Val Glu
  1           5           10           15

CTG CGG TAC GAG ACC GTC GAT CCG TAC GCG GTG CGG CTG ACG TTC CAC
96
Leu Arg Tyr Glu Thr Val Asp Pro Tyr Ala Val Arg Leu Thr Phe His
  20           25           30

CTC CCC GGA GAC GCC CCG GTC ACC TGG GTC TTC GGG CGT GAA CTG CTG
144
Leu Pro Gly Asp Ala Pro Val Thr Trp Val Phe Gly Arg Glu Leu Leu
  35           40           45

GTC GAG GGA GTC CTG GAC GCC GCG GGC GAC GGC GAC GTC CGG GTC TGC
192
Val Glu Gly Val Leu Asp Ala Ala Gly Asp Gly Asp Val Arg Val Cys
  50           55           60

CCG GTG GGG CAG ACG GCC ACC AGG GAG GTG CAC ATC ACC CTC CAG GTC
240
Pro Val Gly Gln Thr Ala Thr Arg Glu Val His Ile Thr Leu Gln Val
  65           70           75           80

GGC TCC GAG CAG GCG CTC TTC CGC GTC GGC AAG GCG CCG CTG CTC GCC
288
Gly Ser Glu Gln Ala Leu Phe Arg Val Gly Lys Ala Pro Leu Leu Ala
  85           90           95

TTC CTC GAC CGC ACC GAC CAG GGC TTG TCG CTC GGC AGC GAG CGG GCA
336
Phe Leu Asp Arg Thr Asp Gln Gly Leu Ser Leu Gly Ser Glu Arg Ala
 100           105           110

CAC GCC GAC TTC GAC AGC CAC CTC GAC GAC GCT CTG AAC CGC AGC CTC
384
His Ala Asp Phe Asp Ser His Leu Asp Asp Ala Leu Asn Arg Ser Leu
 115           120           125

GCC GAG GAG CAG AGC GCC GGC TG
408
Ala Glu Glu Gln Ser Ala Gly

```

130

135

## (2) INFORMATION FOR SEQ ID NO: 5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 135 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```

Met Ser Phe Leu Val Ser Glu Glu Leu Ala Phe Arg Ile Pro Val Glu
 1              5              10              15
Leu Arg Tyr Glu Thr Val Asp Pro Tyr Ala Val Arg Leu Thr Phe His
          20              25              30
Leu Pro Gly Asp Ala Pro Val Thr Trp Val Phe Gly Arg Glu Leu Leu
          35              40              45
Val Glu Gly Val Leu Asp Ala Ala Gly Asp Gly Asp Val Arg Val Cys
          50              55              60
Pro Val Gly Gln Thr Ala Thr Arg Glu Val His Ile Thr Leu Gln Val
          65              70              75              80
Gly Ser Glu Gln Ala Leu Phe Arg Val Gly Lys Ala Pro Leu Leu Ala
          85              90              95
Phe Leu Asp Arg Thr Asp Gln Gly Leu Ser Leu Gly Ser Glu Arg Ala
          100             105             110
His Ala Asp Phe Asp Ser His Leu Asp Asp Ala Leu Asn Arg Ser Leu
          115             120             125
Ala Glu Glu Gln Ser Ala Gly
          130             135

```

## (2) INFORMATION FOR SEQ ID NO: 6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 408 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (iii) HYPOTHETICAL: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Streptomyces goldiniensis
- (B) STRAIN: ATCC 21386

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..408
- (D) OTHER INFORMATION: /product= "SsgA"  
/gene= "ssgA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

ATG AGC TTC CTC GTC TCG GAA GAA CTC TCC TTC CGT ATT CCG GTG GAG  
 48  
 Met Ser Phe Leu Val Ser Glu Glu Leu Ser Phe Arg Ile Pro Val Glu  
 1 5 10 15  
 CTG CGT TAC GAG ACC TGT GAT CCC TAC GCC GTG CGG CTG ACC TTT CAT  
 96  
 Leu Arg Tyr Glu Thr Cys Asp Pro Tyr Ala Val Arg Leu Thr Phe His  
 20 25 30  
 CTG CCC GGA GAT GCC CCG GTG ACC TGG GCG TTC GGG CGG GAG TTG CTC  
 144  
 Leu Pro Gly Asp Ala Pro Val Thr Trp Ala Phe Gly Arg Glu Leu Leu  
 35 40 45  
 ATC GAC GGA GGT CCG CGG CCG TGC GGG GAC GGG GAC GTC CAC ATC GCG  
 192  
 Ile Asp Gly Gly Pro Arg Pro Cys Gly Asp Gly Asp Val His Ile Ala  
 50 55 60  
 CCC GCC GAC CCG GAG ACG TTC GGC GAG GTC CTG ATC CGC CTG CAG GTG  
 240  
 Pro Ala Asp Pro Glu Thr Phe Gly Glu Val Leu Ile Arg Leu Gln Val  
 65 70 75 80  
 GGG AGC GAC CAG GCG ATG TTC CCG GTC GGC ACG GCG CCG CTG GTG GCC  
 288  
 Gly Ser Asp Gln Ala Met Phe Arg Val Gly Thr Ala Pro Leu Val Ala  
 85 90 95  
 TTC CTG GAC CGC ACG GAC AAG ATC GTG CCG CTG GGG CAG GAG CGT TCC  
 336  
 Phe Leu Asp Arg Thr Asp Lys Ile Val Pro Leu Gly Gln Glu Arg Ser  
 100 105 110  
 CTC GCC GAC TTC GAC GCC CTG CTC GAC GAG GCG CTG GAC CGC ATC CTG  
 384  
 Leu Ala Asp Phe Asp Ala Leu Leu Asp Glu Ala Leu Asp Arg Ile Leu  
 115 120 125  
 GCC GAG GAG CAG AAC GCC GGC TG  
 408  
 Ala Glu Glu Gln Asn Ala Gly  
 130 135

## (2) INFORMATION FOR SEQ ID NO: 7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 135 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Met Ser Phe Leu Val Ser Glu Glu Leu Ser Phe Arg Ile Pro Val Glu  
 1 5 10 15  
 Leu Arg Tyr Glu Thr Cys Asp Pro Tyr Ala Val Arg Leu Thr Phe His  
 20 25 30  
 Leu Pro Gly Asp Ala Pro Val Thr Trp Ala Phe Gly Arg Glu Leu Leu  
 35 40 45

Ile Asp Gly Gly Pro Arg Pro Cys Gly Asp Gly Asp Val His Ile Ala  
           50                                  55                                  60  
 Pro Ala Asp Pro Glu Thr Phe Gly Glu Val Leu Ile Arg Leu Gln Val  
       65                                  70                                  75                                  80  
 Gly Ser Asp Gln Ala Met Phe Arg Val Gly Thr Ala Pro Leu Val Ala  
                                   85                                  90                                  95  
 Phe Leu Asp Arg Thr Asp Lys Ile Val Pro Leu Gly Gln Glu Arg Ser  
                                   100                                  105                                  110  
 Leu Ala Asp Phe Asp Ala Leu Leu Asp Glu Ala Leu Asp Arg Ile Leu  
                                   115                                  120                                  125  
 Ala Glu Glu Gln Asn Ala Gly  
       130                                  135

## (2) INFORMATION FOR SEQ ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 408 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Streptomyces netropsis
  - (B) STRAIN: ATCC 23940
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 1..408
  - (D) OTHER INFORMATION: /product= "SsgA"  
                                   /gene= "ssgA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

ATG AGC TTC CTC GTC TCC GAG GAG CTC TCC TTC AAG ATC CCA GTC GAA  
 48  
 Met Ser Phe Leu Val Ser Glu Glu Leu Ser Phe Lys Ile Pro Val Glu  
   1                                  5                                  10                                  15  
 CTG CGA TAC GAG ACC CGG GAT CCC TAC GCG GTG CGG ATG ACC TTC CAC  
 96  
 Leu Arg Tyr Glu Thr Arg Asp Pro Tyr Ala Val Arg Met Thr Phe His  
                                   20                                  25                                  30  
 CTC CCC GGA GAC GCG CCT GTG ACC TGG GCG TTC GGC CGG GAG CTG CTG  
 144  
 Leu Pro Gly Asp Ala Pro Val Thr Trp Ala Phe Gly Arg Glu Leu Leu  
                                   35                                  40                                  45  
 CTC GAC GGG ATC AAC CGC CCG AGC GGC GAC GGC GAC GTC CAC ATC GCC  
 192  
 Leu Asp Gly Ile Asn Arg Pro Ser Gly Asp Gly Asp Val His Ile Ala  
       50                                  55                                  60  
 CCG ACC GAC CCC GAG GGC CTG TCG GAC GTC TCC ATC CGG CTC CAG GTG

240  
 Pro Thr Asp Pro Glu Gly Leu Ser Asp Val Ser Ile Arg Leu Gln Val  
 65 70 75 80

GGC GCG GAC CGC GCC CTC TTC CGT GCA GGC GCC CCG CCG CTG GTC GCC  
 288  
 Gly Ala Asp Arg Ala Leu Phe Arg Ala Gly Ala Pro Pro Leu Val Ala  
 85 90 95

TTC CTC GAC CGC ACG GAC AAG TCG GTG CCG CTC GGT CAG GAA CAG ACT  
 336  
 Phe Leu Asp Arg Thr Asp Lys Ser Val Pro Leu Gly Gln Glu Gln Thr  
 100 105 110

CTG GGT GAC TTC GAG GAC AGC CTG GAG GCC GCG CTC GGC AAG ATC CTC  
 384  
 Leu Gly Asp Phe Glu Asp Ser Leu Glu Ala Ala Leu Gly Lys Ile Leu  
 115 120 125

GCC GAG GAG CAG AAC GCC GGC TG  
 408  
 Ala Glu Glu Gln Asn Ala Gly  
 130 135

## (2) INFORMATION FOR SEQ ID NO: 9:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 135 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

Met Ser Phe Leu Val Ser Glu Glu Leu Ser Phe Lys Ile Pro Val Glu  
 1 5 10 15

Leu Arg Tyr Glu Thr Arg Asp Pro Tyr Ala Val Arg Met Thr Phe His  
 20 25 30

Leu Pro Gly Asp Ala Pro Val Thr Trp Ala Phe Gly Arg Glu Leu Leu  
 35 40 45

Leu Asp Gly Ile Asn Arg Pro Ser Gly Asp Gly Asp Val His Ile Ala  
 50 55 60

Pro Thr Asp Pro Glu Gly Leu Ser Asp Val Ser Ile Arg Leu Gln Val  
 65 70 75 80

Gly Ala Asp Arg Ala Leu Phe Arg Ala Gly Ala Pro Pro Leu Val Ala  
 85 90 95

Phe Leu Asp Arg Thr Asp Lys Ser Val Pro Leu Gly Gln Glu Gln Thr  
 100 105 110

Leu Gly Asp Phe Glu Asp Ser Leu Glu Ala Ala Leu Gly Lys Ile Leu  
 115 120 125

Ala Glu Glu Gln Asn Ala Gly  
 130 135

## (2) INFORMATION FOR SEQ ID NO: 10:



- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 34 nucleotides
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

- (vi) ORIGINAL SOURCE:
  - (C) INDIVIDUAL PRODUCT: ssg1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

GGCGAATTCG AACAGCTACG TGGCGAAGTC GCCA  
34

(2) INFORMATION FOR SEQ ID NO: 11:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 32 nucleotides
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

- (vi) ORIGINAL SOURCE:
  - (C) INDIVIDUAL PRODUCT: ssg2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GTGGGATCCG TGCTCGCGGC GCTGGTCGTC TC  
32

(2) INFORMATION FOR SEQ ID NO: 12:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 32 nucleotides
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

- (vi) ORIGINAL SOURCE:
  - (C) INDIVIDUAL PRODUCT: ssg3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

GGAATTCCA TATGCGCGAG TCGTTCAAG CA  
32

(2) INFORMATION FOR SEQ ID NO: 13:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 nucleotides
  - (B) TYPE: nucleic acid

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL PRODUCT: ssg4

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

CCGGTCAGCC GCGTTCTGC TCCTC

25